Approved by the Academic council
of the Education programme

_____

Protocol No. ___ from __. __. 20

Syllabus
Bayesian statistics
(4 ECTS)

Kyle L. Marquardt, kmarquardt@hse.ru
Department of_____
Meeting Minute # ___dated _____ 20_

1. **Course Description**

a) Prerequisites

"Methodology and research methods of political science" and "Methods of analyzing heterogeneous data"

b) **Abstract**

In this course, students will learn both basic theory of Bayesian statistical analysis as well as important applications that use Bayesian methods. Specifically, the course will focus on using Bayesian methods to analyze survey data and latent concepts, as well as data with a hierarchical structure.

2. **Learning Objectives**

Understanding and applying Bayesian methods with important applications; using these methods to produce innovative research outputs.

3. **Learning Outcomes**

Students will be able to design and interpret Bayesian statistical analyses, using prominent software such as JAGS and STAN. Students will also learn how to design and implement Bayesian multilevel models and measurement models.

4. **Course Plan**

The course will consist of 16 classes, each consisting of a lecture and seminar. Each class will thus be divided into 1) a lecture in which students will be introduced to the theoretical and practical issues involved in a method, and 2) a seminar in which students will work with the instructor and small groups to implement the techniques to which they were exposed.

The class will be divided into six thematic groups, each with one or more lectures: 1) an introduction to Bayesian statistics, 2) Bayesian regression analysis, 3) Bayesian survey analysis, 4) hierarchical models, and 5) latent variable models, and 6) simulation analyses.

More specifically, the plan is as follows:

Class 1 (03.09.19): Introduction to Bayesian statistics 1
Class 2 (10.09.19): Introduction to Bayesian statistics 2: Bayesian mixed methods
Class 3 (17.09.19): Introduction to Bayesian statistics 3: Simulation-based Bayesian analysis
Class 4 (24.09.19): Bayesian regression analysis 1: Linear models
Class 5 (01.10.19): Bayesian regression analysis 2: Reporting results; model choice
Class 6 (08.10.19): Bayesian regression analysis 3: Generalized linear models 1
Class 7 (15.10.19): Bayesian regression analysis 4: Generalized linear models 2
M1 exams (22.10.19): No class
Class 8 (29.10.19): Bayesian survey analysis; missing data
Class 9 (05.11.19): In-class presentations of replication analyses
Class 10 (12.11.19): In-class presentations of replication analyses
Class 11 (19.11.19): Hierarchical models 1: Introduction
Class 12 (26.11.19): Hierarchical models 2: MrP
Class 13 (03.12.19): Latent variable models 1: Measurement models
Class 14 (10.12.19): Latent variable models 2: Common space models; bridging
Class 15 (17.12.19): Latent variable models 3: Dynamic latent variable models
Class 16 (24.12.19): Latent variable models 3: Dynamic latent public opinion models

M2 exams (25-30.12.19): Presentation of final projects

The instructor reserves the option to amend the schedule to provide more in-depth discussions of specific topics as necessary.

5. **Reading List**

c) Required:

Humphreys, M & Jacobs, A.M. (2015). Mixing Methods: A Bayesian Approach. *American Political Science Review*, 109(4), 653–673.

Fariss, C.J. 2014. Respect for Human Rights has Improved Over Time: Modeling the Changing Standard of Accountability. *American Political Science Review*, 108(2), 297–318.

Stegmueller, D., 2013. How many countries for multilevel modeling? A comparison of frequentist and Bayesian approaches. *American Journal of Political Science*, 57(3), pp.748-761.

d) **Optional:**

Treier, S. and Jackman, S., 2008. Democracy as a latent variable. *American Journal of Political Science*, 52(1), pp.201-217.

Pemstein, D., Meserve, S.A. and Melton, J., 2010. Democratic compromise: A latent variable analysis of ten measures of regime type. *Political Analysis*, 18(4), pp.426-449.
Hare, C., Armstrong, D. A., Bakker, R., Carroll, R., & Poole, K. T. (2015). Using Bayesian Aldrich- McKelvey Scaling to Study Citizens' Ideological Preferences and Perceptions. *American Journal of Political Science*, 59(3), 759-774.

6. Grading System

The final grade will be a function of performance on four sets of tasks:

Grade = .2*Participation + .1*Homework + .35*Replication + .35*Final

The tasks are as follows:

1. Participation. Students will attend all lectures and be prepared to participate in seminars discussions. This will account for 1.25%x16=20% of the final grade.

2. Homework. There will be two homework assignments. One assignment will cover basic Bayesian theory, and the second will cover multilevel models. Each assignment will count for 5% of the final grade, for a total of 2x5%=10%.

3. Replication project. Students will replicate a frequentist analysis using a Bayesian framework. This analysis must come from a paper published in a top political science journal in the past five years. Students will present the proposed analysis to the instructor in outline form four weeks prior to the deadline (01.10.2019), though they are strongly encouraged to discuss the project with him in advance. The outline will consist of a brief description of 1) the proposed analysis (with statistical notation) and 2) the data (including source of data).

The final paper (deadline: 29.10.2019) will be no more than 4,000 words long, and will consist of six sections: 1) introduction, 2) data description, 3) model description, 4) results, 5) conclusion, and 6) appendix (appendix does not count toward word limit). In the introduction, students will briefly discuss the theoretical rationale for their analysis. In the data description, students will discuss the data they use and present relevant descriptive statistics (including graphics as needed). In the model description, students will discuss both the frequentist model and the Bayesian model, providing a justification for their modeling choices. In the results section, students will compare and contrast the frequentist and Bayesian results; they will also graphically present the main results from the Bayesian analysis. In the conclusion, students will discuss which model they find most compelling. In the appendix, students will demonstrate that they have conducted necessary diagnostics of the Bayesian model. Students will also provide complete annotated replication code and data.

Students will also provide 10 minute presentations of these papers on either (05.11.19 or 12.11.19), followed by 5 minutes discussant comments from a fellow student; and a 10 minute question and answer period.

For discussant comments, students should replicate their colleagues' analysis before hand using the provided code. They should provide feedback on both the modeling techniques, as well as diagnostics in the appendix.

The replication project will count for 35% of the final grade: 5%*Outline + 15%xPaper + 5%xDefense + 10%xComments.

4) Final project. For the final project, students have two options. First, they can substantially expand upon their replication project by conducting additional analyses that take advantage of the Bayesian framework (e.g. by conducting analyses with theoretically-motivated priors, incorporating a hierarchical structure or latent variable into the analysis), or conduct a new set of analyses with a clear connection to course materials. Again, the project will include an outline (deadline: 19.11.19), paper (deadline: 17.12.19) and an in-class presentation of the results during finals week (with discussant comments that follow the same criteria as those for the replication project).

The final project will count for 35% of the final grade: 5%*Outline + 15%xPaper + 5%xDefense + 10%xComments.

7. Examination Type

The final exam will involve both an oral defense of a final project (presentation and Q&A), as well as discussant comments on another student's project.

8. Methods of Instruction

Lecture and Seminar; small group work and projects

9. Special Equipment and Software Support (if required)

Computer lab w access to internet; R